Lesson 1: Organizing Data

Goals

- Comprehend that a “scatter plot” represents data with two variables and does not represent a function.
- Coordinate (orally and in writing) representations of data in scatter plots and tables.
- Describe (orally and in writing) patterns in representations of data in scatter plots and tables, and use these representations to make predictions.

Learning Targets

- I can organize data to see patterns more clearly.

Lesson Narrative

To open the unit on finding associations between two variables, students are asked to consider a list of unsorted data to notice any patterns or associations. Later, students organize the table to make the pattern more clear and then see a **scatter plot** as a graphic representation of data that can make the association even more obvious. The last activity in the lesson asks students to match different representations of data (MP7) and stresses the importance of labeling each representation so that the meaning is not lost. Scatter plots will be used throughout the unit to help students recognize various associations between variables.

Alignments

Addressing

- 8.SP.A.1: Construct and interpret scatter plots for bivariate measurement data to investigate patterns of association between two quantities. Describe patterns such as clustering, outliers, positive or negative association, linear association, and nonlinear association.

Building Towards

- 8.SP.A: Investigate patterns of association in bivariate data.

Instructional Routines

- MLR5: Co-Craft Questions
- MLR8: Discussion Supports
- Notice and Wonder

Required Materials

Copies of blackline master

Required Preparation

Provide one copy of the blackline master from the activity Tables and Their Scatter Plots for each student.

> ### Student Learning Goals
>
> Let’s find ways to show patterns in data

1.1 Notice and Wonder: Messy Data

Warm Up: 5 minutes

The purpose of this warm-up is to elicit useful ideas for the following activities. While students may notice and wonder many things about the table, the most important discussion points are thinking about any associations or patterns.

Building Towards

- 8.SP.A

Instructional Routines

- Notice and Wonder

Launch

Arrange students in groups of 2. Tell students that they will look at a table. Their job is to think of at least one thing they notice and at least one thing they wonder. Display the table for all to see. Ask students to give a signal when they have noticed or wondered about something. Give students 1 minute of quiet think time and then 1 minute to discuss the things they notice with their partner, followed by a whole-class discussion.

> ### Student Task Statement
>
> Here is a table of data. Each row shows two measurements of a triangle.

length of short side (cm)	length of perimeter (cm)
0.25	1
2	7.5
6.5	22
3	9.5
0.5	2
1.25	3.5
3.5	12.5
1.5	5
4	14
1	2.5

What do you notice? What do you wonder?

Student Response

Answers vary. Sample response:

Notice:

- The perimeter increases when the side length increases.
- The perimeters may be rounded to the nearest half unit.
- The data are not in any obvious order.

Wonder:

- Are the measurements exact?
- Is there a pattern in the data?
- Why measure these triangles?

Activity Synthesis

Ask students to share the things they noticed and wondered. Record and display their responses for all to see. If possible, record the relevant reasoning near the table. After each response, ask the class if they agree or disagree and to explain alternative ways of thinking, referring back to the table each time. If patterns or associations do not come up during the conversation, ask students to discuss these ideas.

1.2 Seeing the Data

10 minutes
In this activity, students continue to look for patterns and associations for variables by sorting the values in a table. They use the table to make predictions about data that is not included in the table (MP2). In the discussion following the activity, students are asked to consider other methods of finding patterns within the data and see a graphical representation that more clearly shows the strong relationship between the variables. The graphical representation can also help improve predictions over those made with only the tables.

Addressing

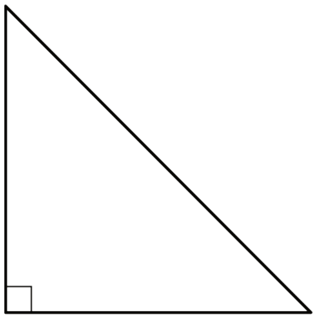
- 8.SP.A.1

Instructional Routines

- MLR8: Discussion Supports

Launch

Keep students in the same groups of 2 from the warm-up. Explain that the data in the warm-up came from fourth grade students who were practicing drawing isosceles right triangles and measuring their perimeters. Display an example of an isosceles right triangle.



Ask students to recall what it means for a triangle to be isosceles and right as well as how to measure the perimeter of a triangle. Ask students if they think there should be a relationship between the length of the two short sides and the entire perimeter. Remind students that specifying 2 sides and the angle between them does determine a unique triangle, so we might expect that knowing the two side lengths and the right angle would be closely related to the length of the perimeter.

Student Task Statement

Here is the table of isosceles right triangle measurements from the warm-up and an empty table.

length of short sides (cm)	length of perimeter (cm)	length of short sides (cm)	length of perimeter (cm)
0.25	1		
2	7.5		
6.5	22		
3	9.5		
0.5	2		
1.25	3.5		
3.5	12.5		
1.5	5		
4	14		
1	2.5		

1. How can you organize the measurements from the first table so that any patterns are easier to see? Write the organized measurements in the empty table.
2. For each of the following lengths, estimate the perimeter of an isosceles right triangle whose short sides have that length. Explain your reasoning for each triangle.
    a. length of short sides is 0.75 cm
    b. length of short sides is 5 cm
    c. length of short sides is 10 cm

Student Response

Answers vary. Sample responses:

1.

length of short sides (cm)	length of perimeter (cm)
0.25	1
0.5	2
1	2.5
1.25	3.5
1.5	5
2	7.5
3	9.5
3.5	12.5
4	14
6.5	22

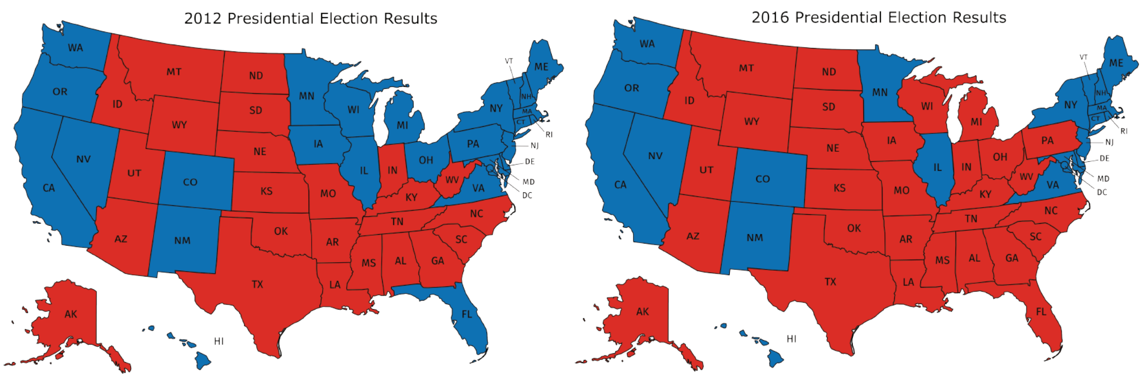
2. a. 2.25 cm. It is between the perimeters for the triangles with short lengths 0.5 cm and 1 cm.

   b. 17 cm. As the length of the short sides increases, so does the perimeter. 17 cm is a little less than halfway between 14 cm and 22 cm, the perimeters for triangles with short sides 4 cm and 6.5 cm.

   c. 30 cm. Each time the length of the short side goes up 1 cm, the perimeter goes up about 3 cm, so when the short length is 10 cm, the perimeter should be about 30 cm.

Are You Ready for More?

In addition to the graphic representations of data you have learned, there are others that make sense in other situations. Examine the maps showing the results of the elections for United States president for 2012 and 2016. In red are the states where a majority of electorate votes were cast for the Republican nominee. In blue are the states where a majority of the electorate votes were cast for the Democrat nominee.



1. What information can you see in these maps that would be more difficult to see in a bar graph showing the number of electorate votes for the 2 main candidates?
2. Why are these representations appropriate for the data that is shown?

Student Response

Answers vary. Sample response:

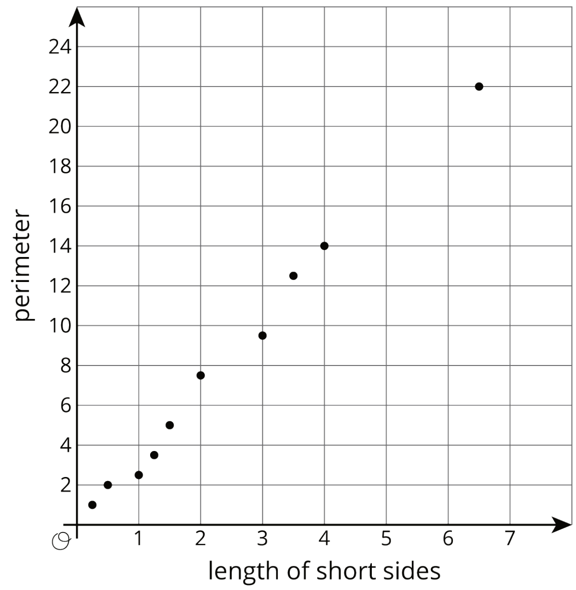
1. The maps show that many of the votes for candidates happen in regions. For example, in the Northeast, they tend to vote for Democrat nominees while in the Southeast, they vote for the Republican nominee.
2. Since electorate votes are cast by state, it makes sense to show the data using a map of the states.

Activity Synthesis

Select students to share their arrangements of the data, the patterns they noticed, and their predictions for the triangles with short side lengths. Display the predictions for all to see.

Ask students to share ideas for other ways to look at the data that might lead to better predictions.

Display a graph of the data for all to see.



To highlight features of the graph, ask:

- "What patterns do you see in the data when it is graphed?" (It is close to linear.)
- "How could you use the graph to estimate the perimeter for the isosceles, right triangles with a short side lengths of 0.75 cm, 5 cm, or 10 cm?" (It is okay for students to not be precise in their answer at this point. Future lessons will give more instruction in how to do this.)

Tell students that this graphical representation of data is called a **scatter plot**. A scatter plot is when two numerical variables are graphed by using one variable as the x-coordinate and the other as the y-coordinate. Data pairs are represented as plotted points.

Note that there is a difference between *time series* graphs and *scatter plots*. In time series graphs, a single variable is recorded at multiple time points and plotted on a graph. In a scatter plot, two variables are measured and plotted on a graph. For scatter plots, time may be one of the variables, but it should be possible to have more than one measurement for the second variable for the same time measurement. For example, when comparing the price of a car to its model year uses the year as one of the variables, but two cars made in the same year could have two different prices. The price for a single car throughout different years would be represented in a time series graph since it only has one price at any given time.

Access for English Language Learners

Speaking: MLR8 Discussion Supports. As students describe the patterns the noticed and their predictions, revoice student ideas to demonstrate mathematical language use by restating a statement as a question (e.g., "As you estimated the perimeters, how did you use the length of the triangles' short side to help you estimate?") in order to clarify, apply appropriate language, and involve more students.
Design Principle(s): Support sense-making; Optimize output (for explanation)

1.3 Tables and Their Scatter Plots

15 minutes
An essential part of creating and understanding scatter plots is interpreting the meaning of the points (MP2). In this activity, students match tables of data with scatter plots representing the same information (MP7). After matching appropriately, students are asked to include titles for the axes of the scatter plots. Following the activity, the importance of the axis labels is discussed.

Addressing

- 8.SP.A.1

Instructional Routines

- MLR5: Co-Craft Questions

Launch

Arrange students in groups of 2. Distribute 1 copy of the tables from the blackline master to each group and resolve any clarifying questions about the data in the tables. In particular, students may wish to know about particular terminology. For example, "kilowatt hours" are a unit of electrical energy that most electricity companies use to measure how much electricity a customer used to determine how much to charge them. Another example is "carat" which is a unit of weight equal to 200 milligrams used for measuring precious stones.

Access for Students with Disabilities

Action and Expression: Internalize Executive Functions. To support development of organizational skills, check in with students within the first 2–3 minutes of work time. Look for students who have matched at least one table with one of the scatter plots accurately. Check to make sure students are in agreement with their matches and either person can justify the decision.
Supports accessibility for: Memory; Organization

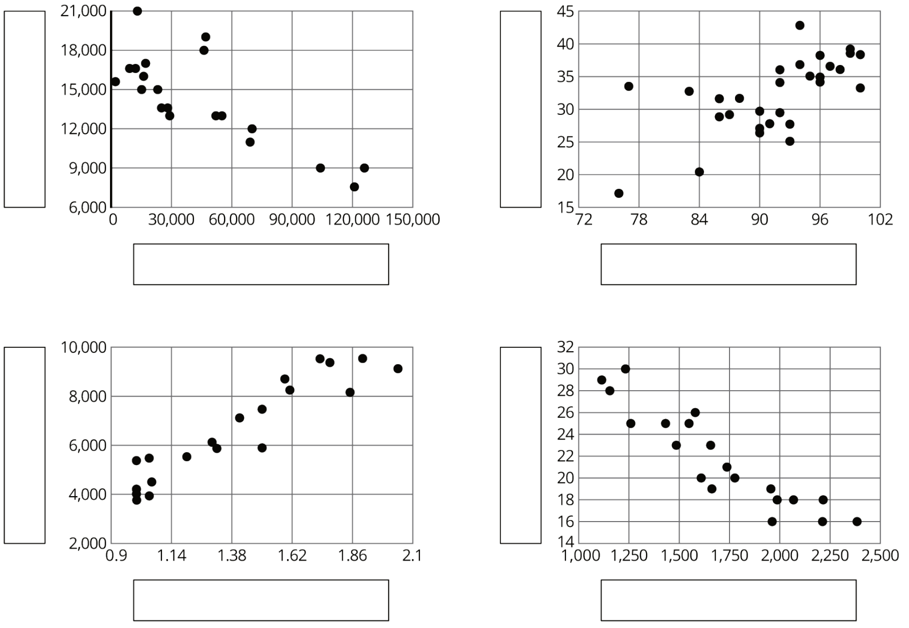
Access for English Language Learners

Conversing, Writing: MLR5 Co-Craft Questions. Display only the four graphs without revealing the questions that follow. Invite pairs of students to write mathematical questions about the graph. Then, invite 2–3 groups to share their questions with the class. Look for questions that ask students to make sense of the relationship between the two quantities represented on each axis. Next, reveal the questions of the activity. This routine allows students to produce the language of mathematical questions and talk about the relationship between quantities that are represented graphically.
Design Principle(s): Maximize meta-awareness; Support sense-making

Student Task Statement

Here are four **scatter plots**. Your teacher will give you four tables of data.

- Match each table with one of the scatter plots.
- Use information from the tables to label the axes for each scatter plot.



Student Response

1. (top left) Title: Used car price vs. mileage, horizontal axis: mileage (miles), vertical axis: price (dollars)
2. (top right) Title: August daily high vs. energy consumed, horizontal axis: daily high (degrees Fahrenheit), vertical axis: energy consumed (kilowatt hours)

3. (bottom left) Title: Diamond prices vs. size in carats, horizontal axis: weight (carats), vertical axis: price (dollars)
4. (bottom right) Title: New car weight vs. fuel efficiency, horizontal axis: mass (kilograms), vertical axis: fuel efficiency (miles per gallon)

Activity Synthesis

The goal of this discussion is for students to develop strategies for relating different representations of data as well as to see the importance of labeling and numbering the axes.

To highlight the axis numbering in preparation for later lessons where students will draw their own scatter plots and need to select a good interval and range, ask:

- "What are some things you notice about the numbers on the axes for the different scatter plots?" (They don't all start at 0 and go up by 1s, 5s, or 10s.)
- "Why do you think it makes sense for the graph comparing temperature and energy to start at 72 and go to 102?" (The minimum temperature is 77 and the maximum temperature is 100.)
- "Why do the numbers for the graph comparing temperature and energy go up by 6 rather than 1 or 10." (The numbers would be too close together if it were 1 and too far apart if they were 10.)

Ask partners to select one of the other axes in the graphs (including vertical axes) and discuss why they may have been chosen the way they were. Select 1–2 groups to share their ideas with the class.

Tell students that they should think about the maximum and minimum values as well as the range (the distance between the maximum and minimum values) when setting the scale for the different axes. For the graphs in this unit, it is not usually essential to include the point $(0, 0)$ in the graph, so that makes the axis labels even more important.

To conclude the discussion, consider asking some of the following questions:

- "How did you match the tables to the scatter plots?"
- "Why is it important to include labels for the axes on scatter plots?" (So that any patterns that are found can be recognized from the scatter plot without having to go back to the table.)
- "Why is it important to include units in the axis labels?" (So that the patterns found can be understood easily)
- "The same data is presented in the tables as in the scatter plots. Which is easier to understand? Explain your reasoning."

Lesson Synthesis

To highlight the progression of representations seen today (unorganized table, ordered table, scatter plot) that help to highlight any patterns that may be present in data, ask:

- "Why would sorting the information in the table be helpful?" (It makes it easier to see the minimum and maximum values of the value you sort by.)
- "When looking for relationships between two variables, what are some graphic representations that might be helpful to use?" (An organized table helps show trends of one variable while another goes from smallest to largest. A scatter plot also shows how one variable changes in relation to the other.)
- "What is a scatter plot and how is it different from plotting points for a function?" (For a function, every input has a single output, but a scatter plot uses data that does not have an input or output and a single value for one variable could have multiple values in the second variable.)

Remind students that, when reorganizing data, it is important to continue to label what the information represents. A table without titles or a scatter plot without labels may show some relationship between numbers, but is meaningless outside of the context.

1.4 Squashed Spheres

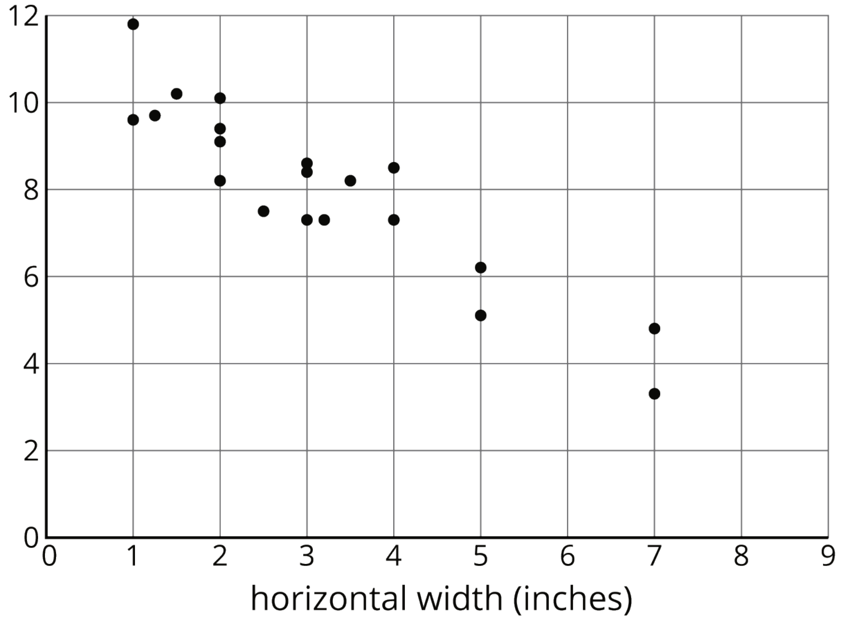
Cool Down: 5 minutes

Addressing

- 8.SP.A.1

Student Task Statement

Twenty rubber spheres were compressed with varying amounts of force. The widths and heights of the resulting shapes were measured. Here is a scatter plot that shows the measurements for each sphere.



1. Label the vertical axis of the scatter plot.

2. If a compressed rubber sphere has a 6-inch width, is its height closer to 5 inches or to 11 inches? Explain your reasoning.

Student Response

1. Vertical height (inches)
2. 5 inches. When the width increases, the height tends to decrease. To keep in line with the rest of the data, the height should be closer to 5 inches than 11 inches.

Student Lesson Summary

Consider the data collected from pulling back a toy car and then letting it go forward. In the first table, the data may not seem to have an obvious pattern. The second table has the same data and shows that both values are increasing together.

Unorganized table

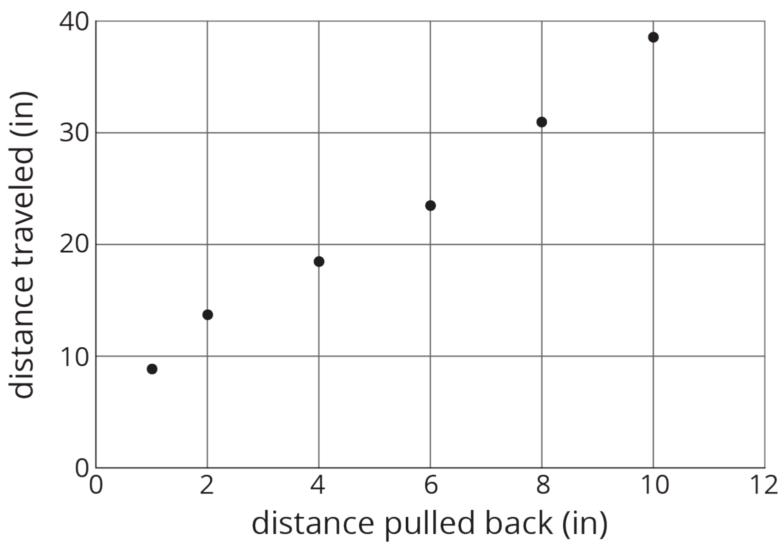
distance pulled back (in)	distance traveled (in)
6	23.57
4	18.48
10	38.66
8	31.12
2	13.86
1	8.95

Organized table

distance pulled back (in)	distance traveled (in)
1	8.95
2	13.86
4	18.48
6	23.57
8	31.12
10	38.66

A scatter plot of the data makes the pattern clear enough that we can estimate how far the car will travel when it is pulled back 5 inches.

Patterns in data can sometimes become more obvious when reorganized in a table or when represented in **scatter plots** or other diagrams. If a pattern is observed, it can sometimes be used to make predictions.



Glossary

- scatter plot

Lesson 1 Practice Problems

Problem 1

Statement

Here is data on the number of cases of whooping cough from 1939 to 1955.

year	number of cases
1941	222,202
1950	120,718
1945	133,792
1942	191,383
1953	37,129
1939	103,188
1951	68,687
1948	74,715
1955	62,786
1952	45,030
1940	183,866
1954	60,866
1944	109,873
1946	109,860
1943	191,890
1949	69,479
1947	156,517

a. Make a new table that orders the data by year.

b. Circle the years in your table that had fewer than 100,000 cases of whooping cough.

c. Based on this data, would you expect 1956 to have closer to 50,000 cases or closer to 100,000 cases?

Solution

a.

year	number of cases
1939	103,188
1940	183,866
1941	222,202
1942	191,383
1943	191,890
1944	109,873
1945	133,792
1946	109,860
1947	156,517
1948	74,715
1949	69,479
1950	120,718
1951	68,687
1952	45,030
1953	37,129
1954	60,886
1955	62,786

b. The years 1948, 1949, 1951, 1952, 1953, 1954, and 1955 had fewer than 100,000 cases of whooping cough.

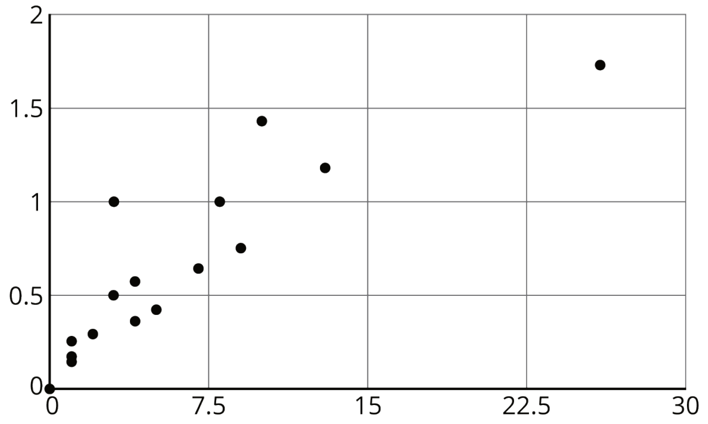
c. This data seems to show the number of cases decreasing over time, so I would expect 1956 to have closer to 50,000 cases than 100,000.

Problem 2

Statement

In volleyball statistics, a block is recorded when a player deflects the ball hit from the opposing team. Additionally, scorekeepers often keep track of the average number of blocks a player records in a game. Here is part of a table that records the number of blocks and blocks per game for each player in a women's volleyball tournament. A scatter plot that goes with the table follows.

blocks	blocks per game
13	1.18
1	0.17
5	0.42
0	0
0	0
7	0.64



Label the axes of the scatter plot with the necessary information.

Solution

The horizontal axis should be labeled "blocks," and the vertical axis should be labeled "blocks per game."

Problem 3

Statement

A cylinder has a radius of 4 cm and a height of 5 cm.

a. What is the volume of the cylinder?

b. What is the volume of the cylinder when its radius is tripled?

c. What is the volume of the cylinder when its radius is halved?

Solution

a. 80π cm^3

b. 720π cm^3

c. 20π cm^3

(From Unit 5, Lesson 18.)